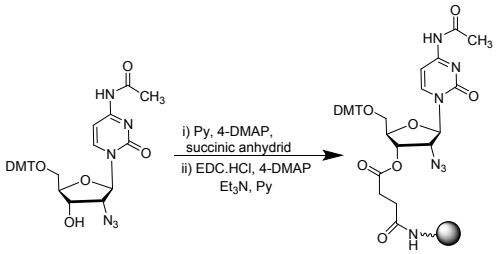
Supporting Information

Reverse transcription through a bulky triazole linkage in RNA: implications for RNA sequencing

Xiong Chen[a], Afaf H. El-Sagheer[b,c] and Tom Brown.[c*]

[a]School of Chemistry, University of Southampton, Highfield, Southampton. SO17 1BJ, UK. [b]Chemistry Branch, Dept. of Science and Mathematics, Faculty of Petroleum and Mining Engineering, Suez University, Suez, 43721, Egypt. [c]Department of Chemistry, University of Oxford, Chemistry Research Laboratory, 12 Mansfield Road, Oxford, OX1 3TA, UK. *e.mail: tom.brown@chem.ox.ac.uk.

Preparation of 5′-*O*-(4,4′-dimethoxytrityl)-2′-azido-2′-deoxycytidine on solid support (2)



1 **2**

The solid support, Amino-SynBase resin 1000/100 (Link Technologies, Glasgow, UK) (1000 Å pore size, loading 59 μmol/g, 0.5 g, 29.5 μmol) was treated with 3% trichloroacetic acid (TCA) in dichloromethane (DCM) for 4 h in a stoppered glass vessel fitted with a sinter and tap. The solvents were then removed by filtration and the support was washed with triethylamine:di-isopropylethylamine (9:1), dichloromethane (DCM) and diethyl ether. The support was dried under vacuum for 1 h then soaked in dry pyridine for 10 min. A solution of succinic anhydride (250 mg, 2.5 mmol) and 4-dimethylaminopyridine (DMAP) (25 mg, 0.21 mmol) in dry pyridine (5 mL) was added and the vessel was rotated for 20 h. The support was washed with pyridine, DCM and diethyl ether, dried and soaked in pyridine for 10 min. A solution of ethyldimethylaminopropylcarbodi-imide hydrochloride (EDC.HCl) (96 mg, 0.5 mmol), DMAP (4 mg, 0.033 mmol) and triethylamine (20 μL) in dry pyridine (5 mL) was added to the solid support in the vessel followed by compound **1** (90 mg, 0.147 mmol)[1] in dry pyridine (2 mL). The reaction vessel was left to rotate for 20 h at room temperature then pentachlorophenol (34 mg, 0.13 mmol) was added and the vessel was rotated for a further 3 h. The solvent was then removed by filtration and the support was washed with pyridine, DCM and diethyl ether. Piperidine (10% in DMF, 10 mL) was added and after rotating the

vessel for 1 min the solid support was washed with DCM and diethyl ether (note that during treatment of the support with piperidine to cap the unreacted succinic acid groups the loading of nucleoside decreased with time due to cleavage of the succinyl linkage. Therefore only a brief 1 min piperidine treatment was carried out).

Capping reagent (oligonucleotide synthesis grade, acetic anhydride/pyridine/tetrahydrofuran: *N*-methyl imidazole in tetrahydrofuran, 1:1, 10 mL, Applied Biosystems) was added and the vessel was rotated for 1 h after which the support was washed with pyridine, DCM and diethyl ether, then left to dry under vacuum overnight. The loading of (**1**) on the support was 25 μmol/g, as determined colorimetrically from the cleaved DMT group.

5′-*O*-(4,4′-dimethoxytrityl)-2′-azido-2′-deoxyuridine was attached to the solid support in the same manner as explained above.

Oligonucleotide Synthesis and purification

General method

Standard DNA phosphoramidites, solid supports, and additional reagents were purchased from Link Technologies, Berry&Associates, Glen Research, Sigma-Aldrich and Applied Biosystems Ltd. All oligonucleotides were synthesized on an Applied Biosystems 394 automated DNA/ RNA synthesizer using a standard 0.2 or 1.0 μmol phosphoramidite cycle of acid-catalyzed detritylation, coupling, capping, and iodine oxidation. Stepwise coupling efficiencies and overall yields were determined by the automated trityl cation conductivity monitoring facility and in all cases were >98.0%. All β-cyanoethyl phosphoramidite monomers were dissolved in anhydrous acetonitrile to a concentration of 0.1 M immediately prior to use. The coupling time for normal A, G, C, and T monomers was 40 s and the coupling time for the 5′-BCN phosphoramidite monomer was extended to 480 s. 2′-azido oligonucleotides were synthesized on the 1.0 μmol scale using the 5′-*O*-(4,4′-dimethoxytrityl)-2′-azido-2′-deoxy(cytidine) or (uridine) solid support (25 μmol/g loading) which were synthesised as explained above. The resin was packed into a twist column (Glen Research) then used to assemble the required oligonucleotide sequence in the 3′- to 5′-direction by standard phosphoramidite oligonucleotide methods. Cleavage of oligonucleotides from the solid support and deprotection was achieved by exposure to concentrated aqueous ammonia solution for 60 min at room temperature followed by heating in a sealed tube for 5 h at 55 °C.

RNA oligonucleotides were prepared using 2′-TBS protected RNA phosphoramidite monomers with *t*-butylphenoxyacetyl protection of the A, G and C nucleobases and unprotected U (Sigma-Aldrich). A solution of 0.3 M benzylthiotetrazole in acetonitrile (Link Technologies) was used as the coupling agent, *t*-butylphenoxyacetic anhydride as the capping agent and 0.1 M iodine as the oxidizing agent (Sigma-Aldrich). All phosphoramidite monomers were dissolved in anhydrous acetonitrile to a concentration of 0.1 M immediately prior to use, and the coupling time for all monomers was 6 min. Stepwise coupling efficiencies were determined by automated trityl cation conductivity monitoring and in all cases were >97%. Cleavage of oligonucleotides from the solid support and deprotection were achieved by exposure to concentrated aqueous ammonia/ethanol (3/1 v/v) for 2 h at room temperature followed by heating in a sealed tube for 45 min at 55 °C.

2′-TBS deprotection of RNA oligoribonucleotides

After cleavage from the solid support and deprotection of the nucleobases and phosphodiesters, oligonucleotides were concentrated to a small volume in *vacuo* (until turbidity starts to appear), transferred to 15 mL plastic tubes and freeze dried. The residue was dissolved in DMSO (300 μL) and triethylamine trihydrofluoride (300 μL) was added after which the reaction mixtures were kept at 65 °C for 2.5 h. Sodium acetate (3 M, 50 μL) and butanol (3 mL) were added with vortexing and the samples were kept at -80 °C for 30 min then centrifuged at 4 °C at 13,000 rpm for 10 min. The supernatant was decanted and the precipitate was washed twice with ethanol (0.75 mL) then dried under vacuum.

Purification of oligonucleotides (DNA or RNA)

The fully deprotected oligonucleotides were purified by reversed-phase HPLC on a Gilson system using a Luna 10μ C8 100Å pore Phenomenex 10x250 mm column with a gradient of acetonitrile in triethylammonium acetate or ammonium acetate (0% to 50% buffer B over 20 min, flow rate 4 mL/min), (buffer A: 0.1 M triethylammonium acetate, pH 7.0, buffer B: 0.1 M triethylammonium acetate, pH 7.0, with 50% acetonitrile). Elution was monitored by UV absorption at 295 nm. After HPLC purification, oligonucleotides were desalted using NAP-25 then NAP-10 columns (GE Healthcare). For RNA oligonucleotides, HPLC using triethylammonium bicarbonate buffer (buffer A: 0.1 M triethylammonium bicarbonate, pH 7.5, buffer B: 0.1 M triethylammonium bicarbonate, pH 7.5, with 50% acetonitrile) was used

to avoid desalting and degradation of the RNA. The fractions from HPLC were evaporated without additional desalting.

All oligonucleotides were characterised by negative-mode electrospray HPLC-mass spectrometry in water, using a Bruker Daltronics micrO-TOF mass spectrometer, using an Acquity UPLC BEH C18 1.7 μm HPLC column (Waters), with a gradient of TEAA/CH_3CN in TEAA/HFIP buffer, increasing from 5-40% buffer B over 14 minutes, with a flow rate of 0.1 mL min^{-1} (buffer A: 10 mM TEAA, 100 mM HFIP (H_2O); buffer B: 20 mM TEAA (CH_3CN)). Raw data was processed/deconvoluted using the DataAnalysis function of the Bruker Daltronics CompassTM 1.3 software package.

Non-templated and template copper catalysed click reactions

A solution of degassed Cu^I click catalyst was prepared from *tris*-hydroxypropyltriazole ligand (THPTA) (2.1 mg in 25 μL water), sodium ascorbate (14 μL of 0.5 M solution in water) and $CuSO_4.5H_2O$ (7 μL of 0.1 M solution in water). In case of the non-templated reaction, this solution was added to a solution of the azide and alkyne oligonucleotides (30 nmol of each in 80 μL water). In case of the templated reaction, a splint oligonucleotide (30 nmol) was added to the mixture of the azide and alkyne oligonucleotides (30 nmol each) in 950 μL of 0.2 M NaCl solution followed by annealing before adding the copper catalyst solution as explained above. The reaction mixture was kept under argon at r.t. for 2 h. Reagents were removed using NAP-10 gel-filtration columns and the ligated RNA was analysed and purified by denaturing 20% polyacrylamide gel electrophoresis (PAGE). The product bands were cut then soaked in Tris-HCl buffer (50 mM Tris-base, 25 mM NaCl, pH 7.5) at 37 °C overnight and desalted by NAP-25 followed by NAP-10 gel-filtration then lyophilized.

Non-templated copper free click reaction

A mixture of the alkyne (5′-BCN) (DNA or RNA) and azide (2′-azide) (RNA) oligonucleotides (30 nmol each) in 60 μL water was kept at room temperature for 2 h then lyophilized before loading to 20% polyacrylamide gel electrophoresis.

The product bands were cut then soaked in Tris-HCl buffer (50 mM Tris-base, 25 mM NaCl, pH 7.5) in case of RNA, or in water in case of DNA at 37 °C overnight, then desalted by NAP-25 followed by NAP-10 gel-filtration and lyophilized.

Reverse transcription of BCN templates

In a total reaction volume of 20 μL, 3 μM primer, 3 μM template, 1x M-MuLV Reverse Transcriptase (RNase H^-) reaction buffer (NEB, 50 mM Tris-HCl, 75 mM KCl, 3 mM $MgCl_2$, pH 8.3 at r.t.), 10 mM DTT (NEB), 0.5 mM dNTP (Promega), 200 u M-MuLV Reverse Transcriptase (1 μL, NEB® #M0253) or M-MuLV Reverse Transcriptase (RNase H^-, 1 μL, NEB® #M0368) were mixed and incubated on a BIO-RAD T100™ Thermal Cycler at 37 °C for different time courses. In some experiments, a specific concentration of Mn^{2+} was added to Mg^{2+} free buffer. In other experiments, Mn^{2+} was added to the buffer in the presence of Mg^{2+}. The enzymatic reaction was stopped by freezing the sample in liquid nitrogen then mixing with an equal volume of formamide and directly loaded onto the gel for PAGE analysis (20 % acrylamide, 600v for 5 h) or diluted to 1 mL and desalted by NAP-10 then freeze-dried overnight for HPLC-MS analysis. To give clear spectra, the RNA template in some HPLC-MS samples was digested by RNase H after reverse-transcription (NEB, 1x MuLV RT (RNase H^-) buffer, 10 mM additional DTT, 5 u RNase H (25 μL) was added to 50 μL reverse-transcription solution, incubated at 37 °C for 6 h, gel-filtered (NAP-10), lyophilized then dissolved in 10 μL water for MS.

To minimize the co-migration of the RNA template and the DNA products during PAGE that results from duplex formation, a 10-fold excess of the full length unlabelled DNA strand complementary to the RNA template was added together with the reverse-transcription product. The mixture was heated at 90 °C for 10 min then cooled down to RT. The samples were then mixed with an equal volume of formamide and loaded for PAGE.

Table 1: oligonucleotide used in this study

1a alkyne and azide ODNs and splints

Code	Sequence (5´ - 3´)	5´	3´	Calc.	Found
K1	r^{P}UUAAUGCUAAUCGUGAUAGGGGU-^{Me}dC^{K}	phosphate	propargyl MedC	7809	7810
Z1	r^{Z}UAGAUCGGAAGAGCGGUUCAG	5'-azido U		6839	6840
K2	r^{P}UUAAUGCUAAUCGUGAUAGGGGUC^{K}	phosphate	propargyl rC	7811	7812
K3	rUUAAUGCUAAUCGUGAUAGGGGUC^{K}		Propagryl rC		
O1	r^{O}CAGAUCGGAAGAGCGGUUCAG	BCN		7040	7040
Z2	r^{P}UUAAUGCUAAUCGUGAUAGGGGUU^{z}	phosphate	2´-azido-2´-dU	7799	7800
Z3	rUUAAUGCUAAUCGUGAUAGGGGUC^{z}		2´-azido-2´-dC	7718	7719
O2	d^{O}CAGATCGGAAGAGCGGTTCAG	BCN		6746	6746
Z4	r^{Z}CAGAUCGGAAGAGCGGUUCAG	5'-azido C		6839	6839
S1	dCCGATCTAGACCCCT			4473	4473
S2	dCCGATCTGGACCCCT			4489	4488

1b. Reverse Transcriptase Templates, sequence in green is DNA

Code	Sequence (5´ - 3´) RNA	Calc.	Found
T1 (K1+Z1)	^{P}UUAAUGCUAAUCGUGAUAGGGGUdC^{Me}$_{t}$UAGAUCGGAAGAGCGGUUCAG	14648	14650
T2 (K3+Z1)	UUAAUGCUAAUCGUGAUAGGGGUC$_{t}$UAGAUCGGAAGAGCGGUUCAG	14570	14571
T3 (Control)	UUAAUGCUAAUCGUGAUAGGGGUCUAGAUCGGAAGAGCGGUUCAG	14570	14572
T4 (K2+Z4)	^{P}UUAAUGCUAAUCGUGAUAGGGGUC$_{t}$CAGAUCGGAAGAGCGGUUCAG	14649	14651
T5 (K1+Z4)	^{P}UUAAUGCUAAUCGUGAUAGGGGUdC^{Me}$_{t}$CAGAUCGGAAGAGCGGUUCAG	14647	14648
T6 (Control)	UUAAUGCUAAUCGUGAUAGGGGUCCAGAUCGGAAGAGCGGUUCAG	14569	14570
T7 (Z2+O1)	^{P}UUAAUGCUAAUCGUGAUAGGGGUU$_{x}$CAGAUCGGAAGAGCGGUUCAG	14839	14839
T8 (Z3+O1)	UUAAUGCUAAUCGUGAUAGGGGUC$_{x}$CAGAUCGGAAGAGCGGUUCAG	14758	14759
T9 (Z3+O2)	UUAAUGCUAAUCGUGAUAGGGGUC$_{x}$CAGATCGGAAGAGCGGTTCAG	14464	14465

1c. Reverse Transcriptase primers

Code	Sequence (5´ - 3´)	position	Calc.	Found
P1	Fam-dCTGAACCGCTCTTC	-7	4711	4711
P2	Fam-dCTGAACCGCTCTTCCGAT	-3	5947	5947
P3	Fam-dCTGAACCGCTCTTCCGATCTA	+0	6853	6854
P4	Fam-dCTGAACCGCTCTTCCGATCTAG	+1	7183	7183
P5	Fam-dCTGAACCGCTCTTCCGATCTAGACC	+4	8074	8075
P6	Fam-dCTGAACCGCTCTTCCGATCTG	+0	6869	6870
P7	Fam-dCTGAACCGCTCTTCCGATCTGG	+1	7199	7199
P8	Fam-dCTGAACCGCTCTTCCGATCTGGACC	+4	8090	8090

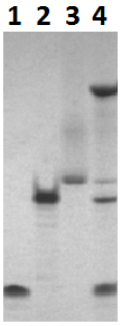
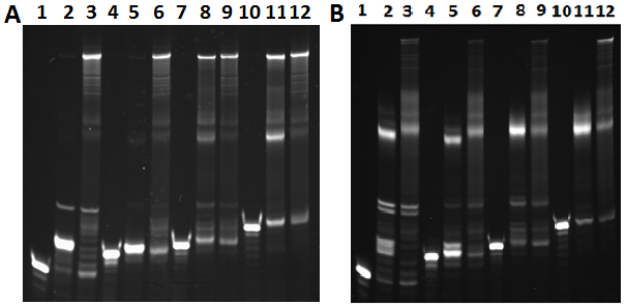


Figure S1: Templated CuAAC click ligation to synthesize T2. Lane 1; S1, lane 2; 5´-azide oligo Z1, lane 3; alkyne oligo K3, lane 4; crude reaction mixture. 20% Polyacrylamide gel.



5´-HO- UUAA UGCU AAUC GUGA UAGG GGUC_tUAGA UCGG AAGA GCGG UUCA G -3´

CCAG ATCT AGCC TTCT CGCC AAGT C-F-5´ P_5

G ATCT AGCC TTCT CGCC AAGT C-F-5´ P_4

ATCT AGCC TTCT CGCC AAGT C-F-5´ P_3

T AGCC TTCT CGCC AAGT C-F-5´ P_2

Figure S2: Mn^{2+} ions enhance the reverse transcriptase read through of the RNA-RNA triazole linkage. Reverse-transcription of triazole template T2 (C_tU) and control T3 by M-MuLV reverse-transcriptase (RNase H^-) in 3mM Mg^{2+} buffer (A) and 3mM Mn^{2+} buffer (B). Lane 1; primer P2 (-3), lane 2 and 3; reverse transcription product of T2 and T3 using primer P2, lane 4; primer P3 (+0), lane 5 and 6; reverse transcription product of T2 and T3 using primer P3, lane 7; primer P4 (+1), lane 8 and 9; reverse transcription product of T2 and T3 using primer P4, lane 10; primer P5 (+4), lane 11 and 12; reverse transcription product of T2 and T3 using primer P5. The transcription products run as two bands (single stranded and double stranded with the RNA template). This problem was overcome in the later gels by adding a 10-fold excess of the full length unlabelled DNA strand complementary to the RNA template to the reverse-transcription product.
P1 (-7) was also used and gave similar results. In some experiments, Mn^{2+} was added to Mg^{2+} containing buffer and results were similar to those which formed from using Mn^{2+} in Mg^{2+} free buffer. This indicates that Mn^{2+} enhances the reverse transcription reaction in the presence or absence of the Mg^{2+} ions.

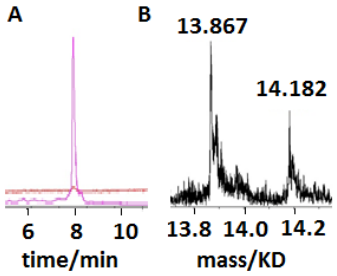


Figure S3: **A) HPLC and B) mass spectrum of the M-MuLV reverse transcriptase (RNase H^-) products (gel purified) of the triazole template T2**: (M minus "G", cal. 13867), (M minus "G" plus "A", cal. 14180). The "plus A" fragment likely resulted from the terminal transferase activity of the reverse-transcriptase. (6 mM Mn^{2+} and primer P1 (-7) were used)

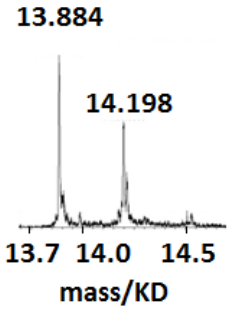


Figure S4: Mass spectrum of reverse transcriptase product of T5: M-MuLV reverse transcriptase with 3 mM $Mg^{2+}$and primer P1 (-7), incubated at 37 °C for 2 h. Calc. Mass: 13883 (M - G); 14196 (M-G+A).

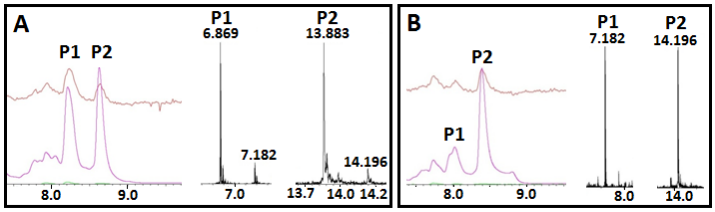


Figure S5: Reverse transcriptase reads through the triazole linkage after overnight incubation. HPLC traces and mass spectra of the reverse transcription product of T8, A) after incubation at 37 °C for 2 h, P1: shows the product formed from RT termination before triazole, Calc. 6869 and RT termination before triazole "+A", Calc. 7183; P2: "M-G" Calc. 13883 and "M-G+A" Calc. 14196. B) After incubation at 37 °C overnight, P1: shows the product formed from RT termination before triazole "+A" Calc. 7183, P2: "M-G+A" Calc. 14196. M-MuLV reverse transcriptase (RNase H^-), 3 mM Mg^{2+} buffer and primer P1 (-7) were used. After the incubation time, digestion of the RNA template by RNase H and desalting by gel-filtration (NAP-10) was carried out.

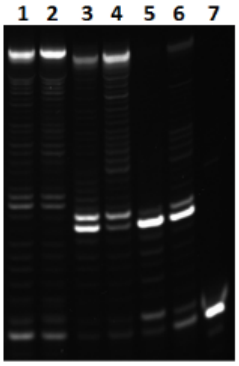


Figure S6: Mn^{2+} ions enhance the reverse transcriptase read through of the RNA-DNA triazole linkage. Lane 1 and 2; reverse transcription of control T6 after 2h and 16h using 3 mM Mn^{2+}, lane 3 and 4; reverse transcription of triazole template T9 after 2h and 16h using 3 mM Mn^{2+}, lane 5 and 6; reverse transcription of triazole template T9 after 2h and 16h using 3 mM Mg^{2+}, lane 7; primer P1 (-7). 20% polyacrylamide gel

DNA sequencing analysis

The reverse transcription products were obtained using unlabelled RT primer (5′-CTGAACCGCTCTTC, same sequence as P1 with no 5′-FAM), templates (T8, C_xC), (T6, control CC) or (T7, U_xC). Mg^{2+} buffer was used in the case of templates T8 and T6 while Mg^{2+} + Mn^{2+} buffer was used in case of template T7.

3 µM primer (300 pmol), 3 µM template (300 pmol), 1x supplied M-MuLV Reverse Transcriptase (RNase H^-) Reaction Buffer (50 mM Tris-HCl, 75 mM KCl, 3 mM $MgCl_2$, or 3 mM $MnCl_2$, pH 8.3 at room temperature), 10 mM DTT, 0.5 mM dNTP (each triphosphate), 1000 U (5 µL) M-MuLV Reverse Transcriptase (RNase H^-)(NEB®) in a total volume of 100 µL. Reverse transcription was carried out at 37 °C for 18h.

cDNAs formed from the reverse transcription of T7 (U_xC), T8 (C_xC) and T6 (CC control) were purified on a 20% polyacrylamide gel (PAGE). Three PCR reactions (50 µL) were carried out using GoTaq DNA polymerase, the PAGE purified DNA and two tailed primers as shown below. The PCR products were purified using a 2% agarose gel followed by extraction using QIAquick Gel Extraction kit (50) Cat. No.28704.

```
                5′-UUAAUGCUAAUCGUGAUAGGGGUCxCAGAUCGGAAGAGCGGUUCAG-3′ CxC template
                3′-AATTACGATTAGCACTATCCCCAG-GTCTAGCCTTCTCGCCAAGTC-5′ RT product
Primers
5′GCATTCGAGCAACGTAAGTTAATGCTAATCGTGA              GCCTTCTCGCCAAGTCGACGGCTGTGTGTATTGG-5′
```

PCR reaction mixture: 25 pmol each primer (final 0.5 µM), 0.002 pmol template, 1x Green GoTaq® Reaction Buffer, 2.5 µL 10 mM dNTP (Promega®, final 0.5 mM), 6.25 u GoTaq® DNA polymerase (5u/µL) in total 50 µL.

PCR condition: 95 °C for 2 min, then 30 cycles of 95 °C for 30 sec, 45 °C for 30 sec, 72 °C for 30 sec.

Cloning and automated Sanger sequencing of these PCR products were performed and the results are shown below. M-MLV RT reads through the unnatural BCN-triazole linkages in RNA with omission of one nucleotide (the first base after the triazole). There are few mutations in other regions of the templates. However similar mutations also appeared in the control sequences suggesting that the mutations far from the triazole site may have occured during the sequencing and cloning process.

T6 CC control, red GG in cDNA strand

```
Sequence 1  GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTGGACCCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 2  GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTGGACCCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 3  GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTGGACCCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 4  GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTGGACCCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 5  GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTGGACCCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 6  GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTGGACCCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 7  GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCG-TCTGGACCCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 8  GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTGGACCCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 9  GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTGGACCCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 10 GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTGGACCCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 11 GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTGGACCCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 12 GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTGGACCCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 13 GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTGGACCCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 14 GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTGGACCCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 15 GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTGGACCCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 16 GGTTATGTGTGTCGGCAGCTGAACCGC-CTTCCGATCTGGACCTCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 17 GGTTATGTGTGTCGGCAGCTGAACCGCTCTTTCGATCTGGACCCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 18 GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTGGACCCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 19 GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTGGACCCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Required    GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTGGACCCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
```

T8 C_xC in RT template should be copied as GG in the cDNA strand

```
Sequence 1  GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTG-ACCCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 2  GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTG-ACCCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 3  GGTTATGTGTGTCGGCAGCTGAACCGTTCTTCCGATCTG-GCCCCTATCACGATTAGC-TTAACTTACGTTGCTCGAATGC
Sequence 4  GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTG-ACCCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 5  GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTG-ACCCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 6  GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTG-ACCCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 7  GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTG-ACCCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 8  GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTG-ACCCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 9  GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTG-A-CCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 10 GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTG-GCCCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 11 GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTG-ACCCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 12 GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTG-ACCCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 13 GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTG-ACCCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 14 GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTG-ACCCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 15 GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTG-ACCCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 16 GGTTATGTGTGTC-GCAGCTGAACCGCTCTTCCGATCTG-ACCCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 17 GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTG-ACCCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Required    GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTGGACCCcTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
```

T7 U_xC in RT template should be copied as GA in the cDNA strand

```
Sequence 1  GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATTTG-ACCCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 2  GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTG-ACCC-TATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 3  GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTG-ACCC-TATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 4  GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTG-ACCCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 5  GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTG-ACCC-TATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 6  GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTG-ACCC-TATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 7  GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTG-ACCC-TATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 8  GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTG-ACCC-TATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 9  GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTG-ACCCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 10 GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTG-ACCCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 11 GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTG-AGCCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 12 GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTG-ACCCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 13 GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTG-ACCC-TATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 14 GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTG-ACCC-TATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 15 GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTG-AC-TATATCACGAGTAGCATTAACTTACGTTGCTCGAATGC
Sequence 16 GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTG-ACCCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 17 GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTG-ACCCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 18 GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTG-ACCCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 19 GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTG-ACCCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 20 GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTG-ACCCCTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
Required    GGTTATGTGTGTCGGCAGCTGAACCGCTCTTCCGATCTGAACCCcTATCACGATTAGCATTAACTTACGTTGCTCGAATGC
```

Figure S7: Sequencing results of clones of the PCR products of the reverse transcripton products of templates T6 (CC control), T8 (C_xC) and (T7 U_xC): From the template sequence (T8) GGGGUCxCAGA, the sequence TCTGGACCCC should be produced but instead of GGA, GA is incorrectly produced. From the template sequence (T7) GGGGUUxCAGA, the sequence TCTGAACCCC should be produced but instead of GAA, GA is incorrectly produced. Therefore we can conclude that the purine base in red is the one that is omitted because the pyrimidine base at the 5'-side of the triazole linkage in the RT template is the one that is not copied (i.e. the one in blue).

Reference

1- K. Fauster, M. Hartl, T. Santner, M. Aigner, C. Kreutz, K. Bister, E. Ennifar and R. Micura, ACS Chem. Biol. 2012, 7, 581-589

2- M. Aigner, M. Hartl, K. Fauster, J. Steger, K. Bister and R. Micura, ChemBioChem, 2011, 12, 47-51.